An Introduction to Data Visualization

Lingrui Zhou*

School of Management, University of San Francisco, San Francisco, United States

*Corresponding author: lzhou31@dons.usfca.edu

Abstract. In the era of information explosion, users are exposed to overwhelming visual stimuli daily. Unprocessed raw data are difficult to understand by the human brain. The business field has more data from more references than ever before. However, the information being presented to most employees, much of the time, is displayed in the form of complicated spreadsheets and hackneyed charts that are missing the crucial insights that lends itself to better business understanding. Thankfully, human beings are very sensitive to the perception of graphics and images; therefore, the imagery behind the numbers in the data graphics makes it easier for people to understand. Data visualization is becoming the best way to communicate with people. The present paper will explore the importance of data visualization and the current uses of data visualization. It will also examine the different tools and potential problems, and a brief on data visualization is included.

Keywords: Data Visualization, Data Visualization tools, Data Analysis.

1. Introduction

Today, information and data are everywhere. Business now has more data from even more sources than ever before. Data that isolated and is kept on computers become invisible over time to users. To be able to have insight into information and understand it, visual interpretations are useful. This turns the once invisible non-actionable data into comprehensible pictures and images. This paper will briefly explore the importance of data visualization. It will also examine some of the uses and problems of operating data visualization.

Data visualization is the graphical representation of the significance of data. The goal of data visualization is to deliver the data in a transparent and compelling way. That is, data visualization's primary purpose is to help in the process of understanding of data. Willam Playfair is credited with first having created and used the area chart and pie chart to display details about the data. Data visualization is not a modern product. Since implementing this concept of data visualization is not new, why is it presently such a prevalent topic?

Data visualization is essential because a visual overview of information is more manageable to identify. It makes it easier to see the trends and patterns of quantitative data. Data visualization is useful for data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modeling output, and presenting results. Visualization can provide a more useful and quicker way to identify trends, patterns, and correlations in the data sets that would remain concealed with just text and numbers.

2. Current Uses of Data Visualiztion

Like many new concepts, technology makes the examination of data and text more accessible. Tools for forging graphical representations of complicated ideas are no longer kept for professionals. Nowadays average individuals are able to have access and learn to use different tools online. With the increase in the amount of data obtainable, it is necessary to be able to interpret increasingly large sets of information, and data visualization is a strong instrument for the job. This is not only required for data analysts; it is critical to understand data visualization in any career. Anyone who works in finance, education, sports management, health monitoring or any other sector needs to visualize data.

2.1. Data visualization in business

At present, we are living in the age of big data. The amount of data used in enterprises, industries, research institutions, and technology development is huge and increasing rapidly every day. Massive volumes of data sets cause it challenging to process at the speed required. Visualization techniques help us to comprehend such complicated data and detect patterns or unusual events.

Data visualization provides a business with vastly improved decision-making because it can enhance the available information and represents it in a pictorial format. In fact, data visualization can help display data from every aspect of the business and other visual elements that are simple for the business to implement. One of the most important decision-making processes in the company's structure should be the process of determining the price. To make the most of any company's earning opportunities, setting the right price for the products is indispensable. To ensure the company determines the right product price, there are many visualization methods that can be used to display business operations and expenses efficiently. For example, if we want to check profit or loss over a year, a great visualization tool to demonstrate that could be a price waterfall visualization. A Waterfall chart is a form of data visualization that displays better insight into the transition of a value over time. Using this particular visualization method, analysts can reveal the effect of changes on a specific value over time. Based on the graph, specific suggestions could be made to review the problem of how to narrow the price-to-profit gap when introducing this information to executives, sales, and product organizations.

Visualization charts with actionable data can display and help to communicate the message effectively. For example, when a company wants to give an instant perspective of how well a business is doing, gives an instant idea of progress on a particular KPI. Containing a simple indicator visualization help to show whether the business is above or below target.

Overall, Data visualization can support organization leaders in identifying patterns and gaps and interpreting the information in a meaningful manner, facilitating a better decision-making process. Nonetheless, effective visualization figures provide an in-depth understanding of the business insight. When financial professionals choose to buy or sell assets, they must track the performance of their investment decisions. Candlestick charts are used as a trading tool to help financial professionals analyze price shifts over time and display important information such as securities, stocks, bond, derivatives, currencies and commodities. By analyzing how the price changes over time, data visualization helps to create patterns, which contribute to better investment strategies. To improve business prospects, data visualization emphasizes the latest trends. At present, we are living in the age of big data, where the amount of data used in enterprises, industries, research institutions, and technology development is huge and growing rapidly every day. Massive volumes of data make it difficult to process at the speed needed. Visualization techniques support us in understanding such complex data and noticing unusual events.

2.2. Data Visualization in Military

For the military, clear and actionable data is critical. Today, the enemy has appeared on the battlefield and posed a threat through digital warfare. It is essential to gather data from various sources, whether organized or unorganized. The amount of data is huge, and the data visualization technology is very necessary to quickly share accurate information in the most concise structure. A better understanding of past data can make more accurate.

2.3. Data Visualization in Education

When it comes to education, one can monitor the students' learning progress throughout the semester, so that the advisers can take prompt actions to help the students who are failing. Educators may visually engage with data. In addition, they can answer students' questions quickly. Therefore, they can make more accurate, data-informed decisions, which eventually share their results with others using instinctive, interactive data visuals. When data is visualized, it is processed quickly by users across all fields.

3. Data Visualization Tools

3.1. Python

Python is considered one of the top-level programming languages for data visualization because it is known for having many libraries that allow for greater flexibility and its large and active scientific computing community. It also controls the specific elements of the created graphics and makes the specifications repeatable through code. Python is also very good at processing data, it provides open-source communities and rich third-party libraries that allow continuous optimization for data visualization.

3.2. R

R is an open-source software environment designed for creating graphics. R is designed for data analysis. Although Python is becoming more and more popular, especially in the areas of machine learning and in-depth learning, the R language still has absolute advantages in data analysis and visualization, with ggplot2 package and its extension package humanized drawing grammar favored by users, especially bioinformatics and medical researchers.

3.3. Power BI

Power BI is able to extract data from a variety of data sources in addition to supporting Microsoft's own products. The drag-and-drop graphical development model used by Power BI will free data analysts from the visual chores and put more effort into data management, algorithm research, and business communication.

3.4. Tableau

Tableau was a commercial intelligence presentation tool for data visualization development and implementation. It could be used for interaction, visual analysis, and dashing applications. It also allowed showing boring data in a simple and friendly diagram. This is one of the most direct and compelling analysis methods. Because not too much technical foundation is needed, anyone could easily learn how to use Tableau and use its visualization tools to deal with and show the data so that they could better analyze the surface data.

4. Concerns Regarding Data Visualization

Although data visualizations can be generated in real-time, it is extremely difficult for a large volume of data to present meaningful and valuable data. Also, the limited screen is not sufficient to display all the required information at the same time. As a result, the visualization is too dense to be useful. When the data is too huge, visualizing every data point can often lead to overlapping and can cause more haze than clearness. Another problem that might drive confusion is the data itself, which is the quality and type of data collected. Attempts to create visualizations with insufficient data are often misleading. Additionally, based on the audience’s past experience or visual knowledge, each user could have different conclusions about the same visualization. Data visualization should be objective and easy to understand, which results in increased accuracy. However, it is important to use the right visualization type for the data at hand to be able to pull correct conclusions from the analysis. Effective visuals display clear data in both presentation and purpose, not distorted in any way.

This is how the human brain works. It can process a significant amount of visual information in parallel at the preconscious level. Unlike our other sensory modalities(sound, touch), our visual system is characterized and suitable for transmitting the information. Therefore, our visual system is the best way to use representation to analyze, comprehend and communicate quantitative data. Visuals can get our attention and enhance memorization and support comprehension is why data visualization is significant. It also enables people to understand the implication of data by summarizing and presenting large amounts of data in detail. However, when visualization is done

wrong, it has the intrinsic limitation of losing information through summarization. Overall, Data visualization makes the data easier to understand for the human mind.

5. Conclusion

Data visualization has a series of attractive characteristics, such as presenting data clearly and effectively. These qualities enable data visualization to penetrate many fields and daily life. It is also an effective tool in elaborating the meaning of information and enabling faster understanding of the data, which eventually leads to ease of design and likely impact on the viewer. Additionally, data visualization can support organizations to adapt and respond to the prevailing market changes and drives them ahead. However, it is important to understand the necessity of building a story supported by data and goals. Further research could be conducted on how to accurately present data with clear and effective views.

References

[1] Munzner, Tamara. *Visualization Analysis and Design*. CRC Press, 2014.

[2] Unwin, A. (2020). Why is Data Visualization Important? What is Important in Data Visualization? . Harvard Data Science Review, 2(1). https://doi.org/10.1162/99608f92.8ae4d525

[3] Schwabish, Jonathan A. "An economist's guide to visualizing data." *Journal of Economic Perspectives* 28.1 (2014): 209-34. Link: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.28.1.209

[4] Steven Batt, Tara Grealis, Oskar Harmon & Paul Tomolonis (2020) Learning Tableau: A data visualization tool, The Journal of Economic Education, 51:3-4, 317-328, DOI: 10.1080/00220485.2020.1804503

[5] Shengjia Cao et al 2021 Research on Python Data Visualization Technology J. Phys.: Conf. Ser. 1757 012122

[6] Rajagopalan, G. (2021). Data Visualization with Python Libraries. In: A Python Data Analyst's Toolkit. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-6399-0_7

[7] *Interactive data visualisation in R - TU Graz*. (n.d.). Retrieved October 7, 2022, from https://courses.isds.tugraz.at/ivis/surveys/ss2018/ivis-ss2018-g5-r-datavis-survey.pdf

[8] Tara J. Brigham (2016) Feast for the Eyes: An Introduction to Data Visualization, Medical Reference Services Quarterly, 35:2, 215-223, DOI: 10.1080/02763869.2016.1152146

[9] D. Zhu, Y. Wang, B. Wei, Z. Guo and F. Wan, "Data Visualization Overview," *2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, 2021, pp. 735-738, doi: 10.1109/ICCASIT53235.2021.9633610.